

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 19:15:44 ; Search time 53 Seconds
(without alignments)
676.779 Million cell updates/sec

Title: US-09-934-455-6

Perfect score: 718

Sequence: 1 MESSNRSSNQSDDKQARF.....FEYLDKVLLELLDSEERK 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	718	100.0	139	10 Q9LTC5	Q9LTC5 arabidopsis
2	355.5	49.5	133	10 P3822	P3822 arabidopsis
3	352.5	49.1	139	10 Q9LTC6	Q9LTC6 arabidopsis
4	351	48.9	131	10 Q9LSX0	Q9LSX0 arabidopsis
5	323	45.0	141	10 Q8H075	Q8H075 oryza sativ
6	293	40.8	243	10 Q8LBU9	Q8LBU9 arabidopsis
7	291.5	40.6	273	10 Q9AQJ3	Q9AQJ3 oryza sativ
8	284.5	39.6	236	10 Q9SUX3	Q9SUX3 arabidopsis
9	284.5	39.6	243	10 Q8VYMO	Q8VYMO arabidopsis
10	283.5	39.5	240	10 Q8H659	Q8H659 lycopersico
11	280.5	39.1	245	10 Q8GZ59	Q8GZ59 lycopersico
12	280	39.0	282	10 Q9M4Y9	Q9M4Y9 mesembryant
13	279	38.9	166	10 Q8S651	Q8S651 oryza sativ
14	277	38.6	303	10 Q949D4	Q949D4 oryza sativ
15	276.5	38.5	277	10 P3332	P3332 nicotiana t
16	276.5	38.5	345	10 Q949D2	Q949D2 oryza sativ

17	273	38.0	244	10 Q9LND1	Q9LND1 arabidopsis
18	272.5	38.0	224	10 Q8LGR8	Q8LGR8 lycopersico
19	272	37.9	226	10 Q22167	Q22167 arabidopsis
20	272	37.9	226	10 Q8L9K1	Q8L9K1 arabidopsis
21	269.5	37.5	233	10 Q40479	Q40479 nicotiana t
22	269.5	37.5	237	10 Q9LW50	Q9LW50 nicotiana s
23	268	37.3	203	10 Q9LDB6	Q9LDB6 catharanthu
24	267	37.2	218	10 Q9ZNR2	Q9ZNR2 arabidopsis
25	265	36.9	291	10 Q40478	Q40478 nicotiana t
26	262	36.5	281	10 Q80387	Q80387 arabidopsis
27	262	36.5	329	10 Q9LVT6	Q9LVT6 arabidopsis
28	261.5	36.4	218	10 Q8LDC8	Q8LDC8 arabidopsis
29	261.5	36.4	221	10 Q8VY90	Q8VY90 arabidopsis
30	261.5	36.4	221	10 Q8LDO4	Q8LDO4 arabidopsis
31	261.5	36.4	250	10 Q9LU55	Q9LU55 arabidopsis
32	261.5	36.4	603	10 Q23591	Q23591 arabidopsis
33	260	36.2	234	10 Q4680	Q4680 lycopersico
34	260	36.2	236	10 Q40476	Q40476 nicotiana t
35	260	36.2	282	10 Q9LW48	Q9LW48 nicotiana s
36	259.5	36.1	213	10 Q943F1	Q943F1 oryza sativ
37	259.5	36.1	282	10 Q8VZ91	Q8VZ91 arabidopsis
38	259	36.1	237	10 Q9FR02	Q9FR02 nicotiana t
39	258.5	36.0	202	10 Q8LXR3	Q8LXR3 glycine max
40	258	35.9	201	10 Q9FKG2	Q9FKG2 arabidopsis
41	258	35.9	210	10 Q9LEM6	Q9LEM6 catharanthu
42	257.5	35.9	214	10 Q9FR33	Q9FR33 lycopersico
43	256.5	35.7	201	10 Q8LDL4	Q8LDL4 arabidopsis
44	256	35.7	241	10 Q8L909	Q8L909 arabidopsis
45	255	35.5	241	10 Q9C5M2	Q9C5M2 arabidopsis

ALIGNMENTS

RESULT 1

Q9LTC5 PRELIMINARY; PRT; 139 AA.
AC Q9LTC5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Nicotiana ERBBP-3-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RT DNA Res. 7:131-135(2000).
DR EMBL; AB025608; BAA95736.1; -.
DR HSSP; O80337; 2GCC.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-Domain; 1.
DR PRINTS; PR00367; ETRHSELENT.
DR ProDom; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.
SQ SEQUENCE 139 AA; 16096 MW; 5B43AB05E93B4050 CRC64;

Query Match 100.0%; Score 718; DB 10; Length 139;

Best Local Similarity 100.0%; Pred. No. 3.9e-68;

Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]


```

Query Match      38.8%; Score 277; DB 10; Length 303;
Best Local Similarity 52.2%; Pred. No. 5.2e-21;
Matches 59; Conservative 14; Mismatches 24; Indels 16; Gaps 1;

QY 20 FGVRRPMPGKFAAEIRDPNRNGARLWLGFTAEAAARAYDRAAFNLRGHLAILNPP-- 77
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 149 YRGVRRPMPGKFAAEIRDPNKGARVWLGFTDSAEAAVAYDRAAYMRGSRALLNPPFLR 208
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 78 -----NRYVPRMDYSLRPPYASSSSSSSSSGSTSTNYSRONORE 116

```

Search completed: December 17, 2003, 19:23:11
Job time : 57 secs

```

Query Match      38.5%; Score 276.5; DB 10; Length 277;
Best Local Similarity 44.9%; Pred. No. 5.2e-21;
Matches 71; Conservative 22; Mismatches 38; Indels 27; Gaps 5;

Qy      2  ESSNRSSNQSDQKOAR---FQVRRRRPWGKFAAIRDPSRNGARLWLGTFTAEAAAR 58
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      71  EYSSKSKVKEIEEKPAKESFQVRRRRPWGKFAAIRDSTRNGRVWLGTFTDPSAAAL 130
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      59  AYDRAAFNLRLGLAILINFPEYYP-----RMDD-----YSLRPPVYASSSSS 99
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      131  AYDQAAFPLMRGTSAILNFPVETVQESLRDMKCHVDECSPVALKKHSLRK--KSLSSK 188
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      100  SSSGGSTSTNVSRQORE---VFEPFYLDKVLBEELDS 134
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      189  KNSSSSSKVRKVKKNVNVVVFEDLGRDYLKOLLSS 226
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 18:28:09 ; Search time 26 Seconds
(without alignments)
514.133 Million cell updates/sec

Title: US-09-934-455-6

Perfect score: 718

Sequence: 1 MESSNRSSNQSDKQARF.....FEYLDKVLLELDSERKR 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	355.5	49.5	133	2	D86175	hypothetical prote
2	293	40.8	243	2	T51989	ethylene responsiv
3	284.5	39.6	236	2	B84718	hypothetical prote
4	276.5	38.5	277	2	T03927	DNA binding protel
5	273	38.0	244	2	B86197	hypothetical prote
6	272.5	38.0	225	2	A85196	EREBP-2 protein [i
7	272.5	38.0	225	2	T51988	ethylene responsiv
8	272	37.9	151	2	T07689	transcription fact
9	272	37.9	226	2	T00409	ethylene-responsiv
10	269.5	37.5	233	2	T02590	DNA binding protel
11	265	36.9	291	2	T02434	DNA binding protel
12	262	36.5	281	2	T52189	ethylene responsiv
13	262	36.5	300	2	T52020	ethylene responsiv
14	261.5	36.4	603	2	E71444	probable EREBP-4 -
15	260	36.2	234	2	T07686	transcription fact
16	260	36.2	236	2	T02432	ethylene-responsiv
17	254	35.4	218	2	F84738	probable AP2 domai
18	247.5	34.5	207	2	T49897	transcription fact
19	245.5	34.2	358	2	D96579	hypothetical prote
20	244.5	34.1	225	2	T52011	ethylene responsiv
21	243.5	33.9	268	2	T04787	hypothetical prote
22	239.5	33.4	222	2	T52019	ethylene responsiv
23	238	33.1	171	2	T00432	probable AP2 domai
24	238	33.1	298	2	T07784	AP2 domain protein
25	237.5	33.1	212	2	T48580	hypothetical prote
26	235.5	32.8	204	2	T49031	cadmium-induced pr
27	235	32.7	303	2	T04541	hypothetical prote
28	234	32.6	315	2	T47955	hypothetical prote
29	228	31.8	262	2	E96747	hypothetical prote

30	227	31.6	245	2	E86168	hypothetical prote
31	225.5	31.4	336	2	B84594	AP2 domain transcr
32	224.5	31.3	225	2	T02433	DNA binding protei
33	224.5	31.3	245	2	B86407	hypothetical prote
34	224	31.2	292	2	T09030	hypothetical prote
35	223	31.1	236	2	T48518	transcription fact
36	220.5	30.7	314	2	E86482	protein F5J5.5 [im
37	216.5	30.2	261	2	B86354	hypothetical prote
38	216.5	30.2	287	2	T01919	probable AP2 domai
39	216	30.1	294	2	C84901	probable AP2 domai
40	215	29.9	328	2	G84826	hypothetical prote
41	214.5	29.9	248	2	T07728	transcription fact
42	214	29.8	261	2	B84610	AP2 domain transcr
43	213.5	29.7	192	2	D96498	RAP2.6 [imported]
44	213.5	29.7	204	2	B86410	protein F3M18.20 [
45	213.5	29.7	263	2	T49870	probable transcrip

ALIGNMENTS

RESULT 1

D86175 hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: D86175

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huijzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D86175

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-133 <STO>

A:Cross-references: GB:AE005172; NID:g1903358; PIDN:AAB70439.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 49.5%; Score 355.5; DB 2; Length 133;

Best Local Similarity 52.2%; Pred. No. 1.1e-27;

Matches 71; Conservative 22; Mismatches 26; Indels 17; Gaps 2;

QY 1 MESSNRSSNQSDKQARFGVRRRPMGKFAAEIRDPNRGARGLWLGTFETAEEAARY 60

Db 1 MQGRRSSGSGGGGAEQKRGVRRRPMGKFAAEIRDPNRGARGLWLGTFETAEDARY 60

QY 61 DRAAFNLRLGLAILNFPNEYPRMDYSLRPPYASSSSSSSGSTSTVNRQNRQVFEF 120

Db 61 DRAAYSMRGKAILNFPHEY-----NMGTGSSSTAANSSSSSQ-QVFEF 103

QY 121 EYLDKVLLELDSSE 136

Db 104 EYLDSDVLDELLEYGE 119

RESULT 2

T51989

ethylene responsive element binding factor 2 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000

C:Accession: T51989

R:Fujimoto, S.Y.; Ohta, M.; Usui, A.; Shinshi, H.; Ohme-Takagi, M.

Plant Cell 12, 393-404, 2000

A:Title: Arabidopsis ethylene responsive element binding factors act as transcriptional

A:Reference number: Z25893

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Accession: A85001; MUID:20083488; PMID:10617198

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-225 <STO>

A:Cross-references: GB:NC_001268; NID:95281024; PIDN:CAB45963.1; GSPDB:GN00140

C:Genetics:

A:Gene: d14785w

A:Map position: 4

Query Match 38.0%; Score 272.5; DB 2; Length 225;
Best Local Similarity 59.8%; Pred. No. 2.9e-19;
Matches 55; Conservative 14; Mismatches 18; Indels 5; Gaps 2;

QY 16 KQARFGRVRRPWGKFAAIRDPSRNGARLWGTFTETAEAAAYDRAAFNLRGHLAIIIN 75

DB 101 KGHYRGVRRPWGKFAAIRDPAKNGARVWLTGTETAEADALAYDRAAFNRGSRALLIN 160

QY 76 FP---NEYYPMDYSLRPPYASSSSSSSGS 104

DB 161 FPLRVNSGEP--DPVRIKRSFSSSNNENGA 190

RESULT 7

T51988

ethylene responsive element binding factor 1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000

C:Accession: T51988

R:Fujimoto, S.Y.; Ohta, M.; Usui, A.; Shinshi, H.; Ohme-Takagi, M.

A:Title: Arabidopsis ethylene responsive element binding factors act as transcriptional

A:Reference number: Z25893

A:Accession: T51988

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-266 <FUJ>

A:Cross-references: EMBL:AB008103; PIDN:BAA32418.1

C:Genetics:

A:Gene: ERF-1

Query Match 38.0%; Score 272.5; DB 2; Length 266;
Best Local Similarity 59.8%; Pred. No. 3.5e-19;
Matches 55; Conservative 14; Mismatches 18; Indels 5; Gaps 2;

QY 16 KQARFGRVRRPWGKFAAIRDPSRNGARLWGTFTETAEAAAYDRAAFNLRGHLAIIIN 75

DB 142 KGHYRGVRRPWGKFAAIRDPAKNGARVWLTGTETAEADALAYDRAAFNRGSRALLIN 201

QY 76 FP---NEYYPMDYSLRPPYASSSSSSSGS 104

DB 202 FPLRVNSGEP--DPVRIKRSFSSSNNENGA 231

RESULT 8

T07689

transcription factor Pti5 - tomato

C:Species: Lycopersicon esculentum (tomato)

C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 11-May-2000

C:Accession: T07689

R:Zhou, J.; Tang, X.; Martin, G.B.

EMBO J. 16, 3207-3218, 1997

A:Title: The Pto kinase conferring resistance to tomato bacterial speck disease interact

A:Reference number: Z16096; MUID:97357308; PMID:9214637

A:Accession: T07689

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-161 <ZHO>

A:Cross-references: EMBL:U09256; NID:g2213782; PIDN:AAC49740.1; PID:g2213783

C:Function:

A:Description: transcription factor

A>Note: binds the GCC box, present in the promoter region of genes encoding pathogenesis

C:Keywords: DNA binding; transcription factor

Query Match 37.9%; Score 272; DB 2; Length 161;
Best Local Similarity 53.8%; Pred. No. 2.2e-19;
Matches 56; Conservative 19; Mismatches 23; Indels 6; Gaps 2;

QY 19 RFRGVRRPWGKFAAIRDPSRNGARLWGTFTETAEAAAYDRAAFNLRGHLAIIINFPN 78

DB 58 KYRGVRRPWGKFAAIRDPSRNGARVWLTGTFTETAEAAAYDRAAFNRGKALLNFP 117

QY 79 EYI---PRMDDYSLRPPYASSSSSSSGSTSTNVSRQNRQREVFE 119

DB 118 EIVNASVSDKLSL---CSNYSITNNNSDSSLNEVSSGTTNDVFE 158

RESULT 9

T00409

ethylene-responsive transcription factor homolog T13E15.15 - Arabidopsis thaliana

N:Alternate names: hypothetical protein At2g44840

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 16-Feb-2001

C:Accession: T00409; D84883

R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Masor

submitted to the EMBL Data Library, July 1997

A:Description: Arabidopsis thaliana chromosome II BAC T13E15 genomic sequence.

A:Reference number: Z14146

A:Accession: T00409

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-226 <ROU>

A:Cross-references: EMBL:AC002388; NID:g3420042; PID:g2344900

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umavam, L.; Tallon, L.C.

euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84883

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-226 <STO>

A:Cross-references: GB:AB002093; NID:g2344900; PIDN:AAC31840.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g44840; T13E15.15

A:Map position: 2

Query Match 37.9%; Score 272; DB 2; Length 226;
Best Local Similarity 48.0%; Pred. No. 3.2e-19;
Matches 61; Conservative 18; Mismatches 30; Indels 18; Gaps 4;

QY 4 SNRSNNQSDDKQARFGRVRRPWGKFAAIRDPSRNGARLWGTFTETAEAAAYDRA 63

DB 76 ATKASGSHAPROKQKQYRGVRRPWGKFAAIRDPKNGARVWLTGTETPDAAAYDRA 135

QY 64 AFNLRGHLAIIINFPN-----EYPRMDDYSLRPPYASSSSSSSGSTSTNVSRQNRQREVFE 118

DB 136 AFQLAGSKAKLNFPHLIGSKYEP---VRIPRRSRPEPSVSDQLTS-----EQKRE-- 184

QY 119 EFELYDD 125

DB 185 --SHVDD 189

RESULT 10

T02590

DNA binding protein EREBP-2 - common tobacco

C:Species: Nicotiana tabacum (common tobacco)

C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 21-Jul-2000

C:Accession: T02590

R:Ohme-Takagi, M.; Shinshi, H.

Plant Cell 7, 173-182, 1995

A:Title: Ethylene-inducible DNA binding proteins that interact with an ethylene responsi

Search completed: December 17, 2003, 19:22:05
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 18:23:03 ; Search time 18 Seconds
(without alignments)
363.151 Million cell updates/sec

Title: US-09-934-455-6

Perfect score: 718

Sequence: 1 MESSNRSSNQSDDKQARF.....FEYLDKVLBLELDSEERK 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	293	40.8	243	1	ERF2_ARATH
2	272.5	38.0	268	1	ERF1_ARATH
3	272	37.9	161	1	PT15_LYCES
4	262	36.5	300	1	ERF5_ARATH
5	244.5	34.1	225	1	ERF3_ARATH
6	240.5	33.5	248	1	AP23_ARATH
7	239.5	33.4	222	1	ERF4_ARATH
8	214.5	29.9	248	1	PT16_LYCES
9	161.5	22.5	164	1	P202_LUPPO
10	116.5	16.2	344	1	RAV1_ARATH
11	116.5	16.2	432	1	AP2_ARATH
12	91	12.7	179	1	PGSG_RAT
13	84.5	11.8	1677	1	VIT_ACITR
14	80	11.1	277	1	YMO5_CABEL
15	78.5	10.9	1244	1	SLA1_YEAST
16	77	10.7	717	1	ARA_DROME
17	75.5	10.5	1337	1	YDM5_SCHPO
18	75	10.4	1759	1	NAH2_SCHPO
19	75	10.4	1532	1	IGA_NERGO
20	74.5	10.4	213	1	VNEN_PAVBO
21	74.5	10.4	955	1	T150_HUMAN
22	74	10.3	708	1	SSRP_MOUSE
23	74	10.3	1823	1	VIT_ICHUN
24	73.5	10.2	506	1	VE2_HPV47
25	73	10.2	843	1	CYF1_BRUMA
26	72.5	10.1	797	1	PAT1_YEAST
27	72.5	10.1	809	1	ENPL_HORVU
28	72.5	10.1	817	1	DLG3_HUMAN
29	72.5	10.1	849	1	DLG3_MOUSE
30	72.5	10.1	849	1	DLG3_RAT
31	72.5	10.1	2004	1	CHDB_HUMAN
32	72	10.0	658	1	CPT2_RAT
33	72	10.0	985	1	4ET_HUMAN

34	71.5	10.0	1912	1	VIT1_CHICK	P87498 gallus gall
35	71	9.9	559	1	FRB7_WHEAT	Q43207 triticum ae
36	70.5	9.8	372	1	YP99_CAEEL	Q09477 caenorhabdi
37	70	9.7	526	1	CLOS_CLOHI	P09870 clostridium
38	70	9.7	983	1	Y144_HUMAN	Q14157 homo sapien
39	69	9.6	1713	1	LMA3_HUMAN	Q16787 homo sapien
40	69	9.6	1901	1	YCF1_TOBAC	P12222 nicotiana t
41	68.5	9.5	486	1	KIC0_XENLA	P02537 xenopus lae
42	68	9.5	619	1	T7L2_HUMAN	Q09q80 homo sapien
43	68	9.5	1162	1	ITAD_HUMAN	Q13349 homo sapien
44	68	9.5	1687	1	VIT2_FUNHE	Q98893 fundulus he
45	67.5	9.4	537	1	YDU2_SCHPO	O13863 schizosacch

ALIGNMENTS

RESULT 1

ID	ERF2_ARATH	STANDARD;	PRT;	243 AA.
AC	O80338;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Ethylene responsive element binding factor 2 (ATERF2).			
GN	ERF2 OR ERF-2 OR AT5G47220 OR MQL5.7.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_taxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20181733; PubMed=10715325;			
RA	Fujimoto S.Y., Ohta M., Usui A., Shinshi H., Ohme-Takagi M.;			
RT	"Arabidopsis ethylene responsive element binding factors act as			
RT	transcriptional activators or repressors of GCC box mediated gene			
RT	expression.";			
RL	Plant Cell 12:393-404(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=20181125; PubMed=10718197;			
RA	Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,			
RA	Tabata S.;			
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence			
RT	features of the regions of 3,076,755 bp covered by sixty P1 and TAC			
RT	clones.";			
RL	DNA Res. 7:31-63(2000).			
CC	-!- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE GCC-BOX			
CC	-!- PATHOGENESIS-RELATED PROMOTER ELEMENT.			
CC	-!- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	-!- SIMILARITY: Contains 1 AP2/ERF domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB008104; BAA32419.1; --			
DR	EMBL; AB018117; BAA97155.1; --			
DR	PIR; T51989; T51989.			
DR	HSSP; O80337; 2GCC.			
DR	TRANSFAC; T04634; --			
DR	InterPro; IPR001471; TF_ERF.			
DR	Pfam; PF00847; AP2-domain; 1.			
DR	ProDom; PD001423; TF_ERF; 1.			
DR	SMART; SM00380; AP2; 1.			
DR	Transcription regulation; DNA-binding; Nuclear protein.			
KW	DOMAIN 23 26			
FT	POLY-GLY.			


```
DR EMBL; AB008103; BAA32418.1; ALT INIT.
DR EMBL; Z97343; CAB45963.1; ALT FRAME.
DR EMBL; AL161546; CAB78753.1; ALT_FRAME.
DR EMBL; AY058174; AAL25588.1; -.
DR EMBL; AY062533; AAL32611.1; -.
DR PDB; 1GCC; 23-MAR-99.
DR PDB; 2GCC; 23-MAR-99.
DR PDB; 3GCC; 23-MAR-99.
DR TRANSFAC; T04633; -.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR PRINTS; PRO0367; ETHRSPLEMT.
DR ProDom; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.
DR Transcription regulation; DNA-binding; Nuclear protein; 3D-structure.
KW Transcription regulation; DNA-binding; Nuclear protein; 3D-structure.
FT DNA_BIND 143 208
FT SEQUENCE 268 AA; 29189 MW; 0FB7839984E97B50 CRC64;
SQ
Query Match 38.0%; Score 272.5; DB 1; Length 268;
Best Local Similarity 59.8%; Pred. No. 4.9e-19;
Matches 55; Conservative 14; Mismatches 18; Indels 5; Gaps 2;
QY 16 KQARPGVRRPQWGFAPAEIRDPNRGRLWGLTFTAEAAAYDRAAFNLRGHLALIN 75
DB 144 KGGYGVRRPQWGFAPAEIRDPNRGRLWGLTFTAEAAAYDRAAFNLRGHLALIN 203
QY 76 FP---NEYVPRDDYSLRPPYASSSSSSSSSGS 104
DB 204 PFLRVNSGEP--DPVRIKSRSSFSSSNENGA 233
RESULT 3
PTIS_LYCES
ID PTIS_LYCES STANDARD; PRT; 161 AA.
AC 004681;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Pathogenesis-related genes transactivator PTIS.
GN PTIS.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97357308; PubMed=9214637;
RA Zhou J., Tang X., Martin G.B.;
RT "The Pto kinase conferring resistance to tomato bacterial speck
RT disease interacts with proteins that bind a cis-element of
RT pathogenesis-related genes."
RL EMBO J. 16:3207-3218(1997).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE GCC-BOX
CC -!- PATHOGENESIS-RELATED PROMOTER ELEMENT. ACTIVATES PLANTS DEFENSE
CC GENES.
CC -!- SUBUNIT: INTERACTS WITH THE PTO KINASE.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U89256; AAC49740.1; -.
CC PIR; T07689; T07689.
CC HSP; O80337; 2GCC.
CC TRANSFAC; T04716; -.
CC InterPro; IPR001471; TF_ERF.
KW Transcription regulation; DNA-binding; Nuclear protein.

DR Pfam; PF00847; AP2-domain; 1.
DR PRINTS; PRO0367; ETHRSPLEMT.
DR ProDom; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.
DR Transcription regulation; DNA-binding; Activator; Nuclear protein.
FT DNA_BIND 56 121
FT SEQUENCE 161 AA; 18051 MW; 67F4D2D4C0B37FB0 CRC64;
SQ
Query Match 37.9%; Score 272; DB 1; Length 161;
Best Local Similarity 53.8%; Pred. No. 2.9e-19;
Matches 56; Conservative 19; Mismatches 23; Indels 6; Gaps 2;
QY 19 RFRGVRPRPQWGFAPAEIRDPNRGRLWGLTFTAEAAAYDRAAFNLRGHLALIN 78
DB 58 KYRGVRRPQWGFAPAEIRDPNRGRLWGLTFTAEAAAYDRAAFNLRGHLALIN 117
QY 79 EYV---PRMDDYSLRPPYASSSSSSSSSGS 119
DB 118 EIVNASVSDKLSL---CSNYYTTNNNSDSSLNVSSTNDVFE 158
RESULT 4
ERFS_ARATH
ID ERFS_ARATH STANDARD; PRT; 300 AA.
AC 080341;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Ethylene responsive element binding factor 5 (ATERF5).
DE Ethylene responsive element binding factor 5 (ATERF5).
GN ERFS OR ERP-5 OR AT5G47230 OR MQL5.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181733; PubMed=10715325;
RA Fujimoto S.Y., Ohta M., Usui A., Shinsui H., Ohme-Takagi M.;
RT "Arabidopsis ethylene responsive element binding factors act as
RT transcriptional activators or repressors of GCC box mediated gene
RT expression."
RL Plant Cell 12:393-404(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGECC).";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE GCC-BOX
CC -!- PATHOGENESIS-RELATED PROMOTER ELEMENT.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB008107; BAA32422.1; -.
CC EMBL; AF385709; AAK60301.1; -.
CC PIR; T52020; T52020.
CC HSP; O80337; 2GCC.
CC TRANSFAC; T04637; -.
CC InterPro; IPR001471; TF_ERF.
CC Pfam; PF00847; AP2-domain; 1.
CC ProDom; PD001423; TF_ERF; 1.
CC SMART; SM00380; AP2; 1.
CC Transcription regulation; DNA-binding; Nuclear protein.
```

```
FT DOMAIN 51 55 POLY-SER.
SQ DNA BIND 153 218 AP2/ERF.
SEQUENCE 300 AA; 33810 MW; 1189D46A28F7251 CRC64;

Query Match 36.5%; Score 262; DB 1; Length 300;
Best Local Similarity 65.4%; Pred. No. 5.7e-18;
Matches 51; Conservative 13; Mismatches 10; Indels 4; Gaps 2;

QY 12 SODKQARFGRVRRPWGKFAAEIRDPNRGRLWLTFTTAEAAAYDRAAFNLRGHL 71
Db 149 SEEEK-HYRGVQRPWGKFAAEIRDPNRGRLWLTFTTAEAAAYDRAAFNLRGHL 207

QY 72 AILNFPNE---YYPRMD 86
Db 208 AILNFPLEVGKWKPRADE 225

RESULT 5
ERF3 ARATH STANDARD; PRT; 225 AA.
AC O80339;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ethylene responsive element binding factor 3 (ATERF3).
GN ERF3 OR ERF-3 OR ATLG50640 OR F17J6.16 OR F11F12.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181733; PubMed=10715325;
RA Fujimoto S.Y., Ohta M., Usui A., Shinshi H., Ohme-Takagi M.;
RT "Arabidopsis ethylene responsive element binding factors act as
transcriptional activators or repressors of GCC box mediated gene
expression.";
RL Plant Cell 12:393-404(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Bucher E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.P., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820(2000).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE GCC-BOX
CC -1- PATHOGENESIS-RELATED PROMOTER ELEMENT.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 1 AP2/ERF domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/)
CC -----

or send an email to license@isb-sib.ch.
CC -----
CC EMBL; AB008105; BAA32420.1; -
CC EMBL; AC079279; AAG51201.1; -
DR EMBL; AC012561; AAF87871.1; -
DR PIR; T52011; T52011.
DR HSSP; O80337; 2GCC.
DR TRANSFAC; T04635; -
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR PRINTS; PRO0367; ETHRSPPELEMNT.
DR ProDom; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 25 89 AP2/ERF.
FT DOMAIN 90 93 POLY-PRO.
FT DOMAIN 175 178 POLY-SER.
FT DOMAIN 181 185 POLY-ASP.
SQ SEQUENCE 225 AA; 25216 MW; C709BEABF306776B CRC64;

Query Match 34.1%; Score 244.5; DB 1; Length 225;
Best Local Similarity 47.1%; Pred. No. 1.9e-16;
Matches 57; Conservative 13; Mismatches 24; Indels 27; Gaps 3;

QY 16 KQARFGRVRRPWGKFAAEIRDPNRGRLWLTFTTAEAAAYDRAAFNLRGHLAIN 75
Db 24 KEIRFGVRRPWGKFAAEIRDPWK-ARVWLTGTFDSAEAAAYDRAARNLRGPKAKTN 82

QY 76 FP-----NEYPRMD-----DYSLRPPYASSSSSSSGSTSTNV 109
Db 83 FPDSSSSPPPNLRFNRQNRQNVDPFMDHRLFTDHOQQFPVNRPTSSSSMSTVESF 142

QY 110 S 110
Db 143 S 143

RESULT 6
AP23 ARATH STANDARD; PRT; 248 AA.
AC P42736; O23105;
DT 01-NOV-1995 (Rel. 32, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE AP2 domain transcription factor RAP2.3 (Related to AP2 protein 3)
DE (Cadmium-induced protein AS30).
GN AT3G16770 OR MGL6.1 OR MGL6.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX Choi S., Baek E., Lee S.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97338147; PubMed=9192694;
RA Okamoto J.K., Caster B., Villarroel R., van Montagu M., Jofuku K.D.;
RT "The AP2 domain of APTALA2 defines a large new family of DNA binding
proteins in Arabidopsis.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:7076-7081(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
features of the regions of 4,504,864 bp covered by sixty P1 and TAC
clones.";
RL DNA Res. 7:131-135(2000).
```



```

RN RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAPFs) sequenced by the
RL SSP consortium (Salk/Stanford/PGEC).";
RN [5] Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- INDUCTION: By cadmium.
CC -1- SIMILARITY: Contains 1 AP2/ERF domain.
CC -1- CAUTION: Ref.1 sequence is most probably a cloning artifact and
CC differs from that shown due to an insertion into the sequence.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z37504; CAA95734.1; ALT_SEQ.
DR EMBL; A7003096; AAC49769.1; -.
DR EMBL; AB022217; BAB02769.1; -.
DR EMBL; AY059917; AAL24399.1; -.
DR EMBL; AY035100; AAK59605.1; -.
DR EMBL; AV142562; AAN13131.1; -.
DR EMBL; AV087488; AAM65031.1; -.
DR HSP; O80337; 2GCC.
DR TRANSFAC; T02658; -.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR PRINTS; PR00367; ETHRSPELEMT.
DR ProDom; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.
DR Transcription regulation; DNA-binding; Nuclear protein; Cadmium.
FT DNA_BIND 76 141
FT DOMAIN 140 143 POLY-PRO.
SQ SEQUENCE 248 AA; 27758 MW; 317D50CC514782F1 CRC64;

Query Match 33.5%; Score 240.5; DB 1; Length 248;
Best Local Similarity 51.0%; Pred. No. 5.2e-16;
Matches .50; Conservative 17; Mismatches 28; Indels 3; Gaps 2;

QY 20 FRCVRRPWGKFAAEIRDPSSRCARLWLGTFETAEAAAYDPAAFNLRGHLAILNFPNE 79
DB 79 YGIRKRPWGWAAEIRDP-RKGRVWVLGTFNTAEAAAYDVAAKQIRGDKAKLNFDDL 137

QY 80 YPRMDYSLRPYASSSSSGSTSTNVSQNQREV 117
DB 138 HHPPPNYT--PPSPSRSTDQPPAKKVCVVSQSESEL 173

RESULT 7
ERF4 ARATH
ID ERF4 ARATH STANDARD; PRT; 222 AA.
AC O80337;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ethylene responsive element binding factor 4 (AtERF4).
GN ERF4 OR ERF-4 OR AT3G5210 OR K7L4.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;

```

```

RN RP SEQUENCE FROM N.A.
RX MEDLINE=20181733; PubMed=10715325;
RA Fujimoto S.Y., Ohta M., Usui A., Shinshi H., Ohme-Takagi M.;
RT "Arabidopsis ethylene responsive element binding factors act as
RT transcriptional activators or repressors of GCC box mediated gene
RT expression.";
RL Plant Cell 12:393-404(2000).
RN [2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unsel M.,
RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA De Simone V., Choisine N., Artiguenave F., Robert C., Bottier P.,
RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Queirer F.,
RA Schaefer P., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Erle H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,
RA Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Liauro C., Fumelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Agirion A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujil C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Millitscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana".
RL Nature 408:820-822(2000).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE GCC-BOX
CC PATHOGENESIS-RELATED PROMOTER ELEMENT.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 1 AP2/ERF domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB008106; BAA32421.1; -.
DR EMBL; AP000413; BAB02150.1; -.
DR PIR; T52019; T52019.
DR HSP; O80337; 2GCC.
DR TRANSFAC; T04636; -.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR PRINTS; PR00367; ETHRSPELEMT.
DR ProDom; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.
DR Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 22 86
FT DOMAIN 132 137 POLY-GLY.
SQ SEQUENCE 222 AA; 23741 MW; CEF16ABA5ACFB073 CRC64;

Query Match 33.4%; Score 239.5; DB 1; Length 222;
Best Local Similarity 44.3%; Pred. No. 5.7e-16;
Matches 54; Conservative 20; Mismatches 29; Indels 19; Gaps 4;

```

```
Qy 3 SSNRSSNNQSDDKQARFRGVRPPWKGFAAEIRDPNRNGARLWLGTFETAEEAARAYDR 62
Db 12 TTQTHNNA-----KEIRYGRVKRPGWGRYAAEIRDPGKK-TRVWLGTFTAEAAAYDT 66
Qy 63 AAFNLRGHLAALNPNYPRMDYSLRPPYASSSSSSSS-----GSTSTNVS 110
Db 67 AARDFRGAKAKTNFT--FUELSQKQVPTGFARSPQSSTLDCASPTLVVPVSATAGNVP 124
Qy 111 RQ 112
Db 125 PQ 126

RESULT 8
PT16_LYCES
ID PT16_LYCES STANDARD; PRT; 248 AA.
AC 004682;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pathogenesis-related genes transcriptional activator PT16.
GN PT16.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97357308; PubMed=9214637;
RA Zhou J., Tang X., Martin G.B.;
RT "The Pto kinase conferring resistance to tomato bacterial speck
RT disease interacts with proteins that bind a cis-element of
RT pathogenesis-related genes.";
RL EMO J. 16:3207-3218 (1997).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE GCC-BOX
CC -!- PATHOGENESIS-RELATED PROMOTER ELEMENT. ACTIVATES PLANTS DEFENSE
CC GENES.
CC -!- SUBUNIT: INTERACTS WITH THE PTO KINASE.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U89257; AAC49741.1; -
CC PIR; T07728; T07728.
CC HSP; O80337; 2GCC.
CC TRANSFAC; T04717; -.
CC InterPro; IPR001471; TF ERF.
CC Pfam; PF00847; AP2-domain; 1.
CC ProDom; PD001423; TF ERF; 1.
CC SMART; SM00380; AP2_1.
CC Transcription regulation; DNA-binding; Activator; Nuclear protein.
CC DNA_BIND 95 159 AP2/ERF.
CC SEQUENCE 248 AA; 27913 MW; 115BD4DDEB1B992 CRC64;

Query Match 29.9%; Score 214.5; DB 1; Length 248;
Best Local Similarity 53.9%; Pred. No. 1.6e-13;
Matches 41; Conservative 15; Mismatches 19; Indels 1; Gaps 1;

Qy 2 SSNRSSNNQSDDKQARFRGVRPPWKGFAAEIRDPNRNGARLWLGTFETAEEAARAYD 61
Db 80 DRKRSVSPSDVTRKKFGVRQRPWGRWAAEIRDPTR-GKRVWLGTFTPEAAVVD 138
Qy 62 RAAPNLRGHLAALNPNP 77
Db 139 KAAVKLKGPDVATNFP 154

Qy 3 SSNRSSNNQSDDKQARFRGVRPPWKGFAAEIRDPNRNGARLWLGTFETAEEAARAYDR 62
Db 12 TTQTHNNA-----KEIRYGRVKRPGWGRYAAEIRDPGKK-TRVWLGTFTAEAAAYDT 66
Qy 63 AAFNLRGHLAALNPNYPRMDYSLRPPYASSSSSSSS-----GSTSTNVS 110
Db 67 AARDFRGAKAKTNFT--FUELSQKQVPTGFARSPQSSTLDCASPTLVVPVSATAGNVP 124
Qy 111 RQ 112
Db 125 PQ 126

RESULT 9
PZ02_LUPPO
ID PZ02_LUPPO STANDARD; PRT; 164 AA.
AC P16146;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PZ02 protein.
GN PZ02.
OS Lupinus polyphyllus (large-leaved lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genistaceae; Lupinus.
OX NCBI_TaxID=3874;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91355854; PubMed=2103434;
RA Perrey R., Warskulat U., Wink M.;
RT "Molecular cloning of a lupin-specific gene from a cDNA library of
RT suspension-cultured cells of Lupinus polyphyllus.";
RL Plant Mol. Biol. 15:175-176 (1990).
CC -!- FUNCTION: ESSENTIAL FOR ALL LUPIN CELLS INDEPENDENT OF THE
CC RESPECTIVE TISSUE.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X51767; CAA36069.1; -
CC PIR; S11881; S11881.
CC HSP; O80337; 2GCC.
CC InterPro; IPR001471; TF ERF.
CC Pfam; PF00847; AP2-domain; 1.
CC PRINTS; PR00367; ETRSPLEMT.
CC ProDom; PD001423; TF ERF; 1.
CC SMART; SM00380; AP2_1.
CC Transcription regulation; DNA-binding; Nuclear protein.
CC DNA_BIND 4 69 AP2/ERF.
CC SEQUENCE 164 AA; 18606 MW; 87BF64C0122ECB00 CRC64;

Query Match 22.5%; Score 161.5; DB 1; Length 164;
Best Local Similarity 48.8%; Pred. No. 1.2e-08;
Matches 41; Conservative 9; Mismatches 21; Indels 13; Gaps 3;

Qy 17 QARFRGVRPPWKGFAAEIRDPNRNGARLWLGTFETAEEAARAYDRAAFNLRGHLAALNF 76
Db 5 QQRYGRGVRQHWGWSVSEIRH-SILKTRIQGTFTSAEDARAYDEARLMCGTRATNF 63
Qy 77 PNEYPRMDYSLRPPYASSSSSS 100
Db 64 P-----YN---PNASQSSSS 75

RESULT 10
RAV1_ARATH
ID RAV1_ARATH STANDARD; PRT; 344 AA.
AC Q9ZWM9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA-binding protein RAV1.
GN RAV1 OR ATG13260 OR T6J4.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
```

OC eucoside II; Brassicales; Brassicaceae; Arabidopsis.
 RN NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN=ecv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA White O., Alonso J., Altai H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egtu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ecv. Columbia;
 RX Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RA "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SSP consortium (Salk/Stanford/PGECC).";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Binds specifically to bipartite recognition sequences
 CC composed of two unrelated motifs, 5'-CAACA-3' and 5'-CACCTG-3'.
 CC -!- SUBUNIT: Monomer.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- TISSUE SPECIFICITY: Expressed in all tissues examined: Roots,
 CC rosette leaves, cauline leaves, inflorescence stems, flowers and
 CC siliques. Highest expression in roots and rosette leaves. Very low
 CC expression in flowers.
 CC -!- DOMAIN: Contains two distinct DNA-binding domains. One is located
 CC in the N-terminal region and binds to the 5'-CAACA-3' motif. The
 CC second is located in the C-terminal region and binds to the
 CC 5'-CACCTG-3' motif.
 CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; AB013886; BAA34250.1; -;
 DR EMBL; AC011810; AAG09554.1; -;
 DR EMBL; AY063855; AAL36211.1; -;
 DR EMBL; AY091291; AAM14230.1; -;
 DR F01; T51329; T51329;
 DR HSSP; O80337; 2GCC.
 DR InterPro; IPR003340; TF.B3.
 DR InterPro; IPR001471; TF.ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR Pfam; PF02362; B3; 1.

DR PRINTS; PRO0367; ETRSPLEMMNT.
 DR ProDom; PD001423; TF.ERF; 1.
 KW DNA-binding; Nuclear protein.
 FT DNA_BIND 59 121 AP2/ERF.
 FT DOMAIN 145 338 B3.
 SQ SEQUENCE 344 AA; 38597 MW; 7349B640B3505823 CRC64;
 Query Match 16.2%; Score 116.5; DB 1; Length 344;
 Best Local Similarity 27.9%; Pred. No. 0.00061;
 Matches 36; Conservative 26; Mismatches 42; Indels 25; Gaps 5;
 QY 11 OSODDKQARFRGVRRPWGKFAAEIRDPNRGALWLGCTFAEAAAYDAAFNLRGH 70
 DB 53 ESRKLPSSKYGVVQPNRGWGAQIYERHQ---RVVLGTENEDEAAAYDAVHRFR 109
 QY 71 LAIINFNPNEYPRMDYSLRPPYASSSSSSSGSTSTNVSRQNRQFEVFEYLDKVL 130
 DB 110 DAVTNFKD---VKMDEDEV--DFLNSHKS-----EIVDM--LRKHTYNE 147
 QY 131 LLDSEERKR 139
 DB 148 ELEQSKRRR 156
 RESULT 11
 ID AP2_ARATH STANDARD; PRT; 432 AA.
 AC P47327;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Floral homeotic protein APETALA2.
 GN AP2 OR AT4G36920 OR AP22.49 OR C7A10.440.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A., AND MUTANTS AP2-1 AND AP2-5.
 RC STRAIN=ecv. Columbia; TISSUE=Flower;
 RX MEDLINE=95003709; PubMed=7919989;
 RA Jofuku K.D., den Boer B.G.W., van Montagu M., Okamura J.K.;
 RT "Control of Arabidopsis flower and seed development by the homeotic
 RT gene APETALA2.";
 RL Plant Cell 6:1211-1225(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ecv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 RA Kreis M., Delsen T., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grynoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weizenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirks W.,
 RA Moolman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysschaert C., Giesen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Grandrath K., Dauner D., Herzl A.,
 RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Quigley F., Ciabaud G., Muendlein A., Felber R.,

SMART; SMO0380; AP2; 2.
Flowering; Transcription regulation; Activator; Developmental protein;
Nuclear protein; DNA-binding; Repeat.
DOMAIN 14 50
FT FT
DOMAIN 119 128
DNA_BIND 128 192
DNA_BIND 128 192
DNA_BIND 220 286
MUTAGEN 159 159
MUTAGEN 251 251
MUTAGEN 420 420
SEQUENCE 432 AA; 47833 MW; 376A35529IEFFB42 CRG64;
Query Match 16.2%; Score 116.5; DB 1; Length 432;
Best Local Similarity 37.5%; Pred. No. 0.0008;
Matches 33; Conservative 16; Mismatches 24; Indels 15; Gaps 4;
QY 1 MESSNRNNQSODDKQAPRGRV-RRPGKFAAEIRDSRCARLWLGTFETAESAARA 59
DB 118 LKSSRGPRSR-----SQYRGVTYVRTGWESHIWD---CGKVYLGGDFATAHAARA 169
QY 60 YDRAAFNLRLGHAILNFPNEYPRMDDY 87
DB 170 YDRAAIKFRGVADINF-----NIDDDY 191
RESULT 12
PGSG RAT STANDARD; PRT; 179 AA.
ID PGSG RAT
AC P04917;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Secretary granule proteoglycan core protein precursor (Chondroitin
DE sulfate proteoglycan core protein) (Proteoglycan 10K core protein)
DE (PGI9 core protein) (Cytolytic granule proteoglycan core protein).
DE PRG1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=88213412; PubMed=3366780;
RA Avraham S., Stevens R.L., Gartner M.C., Austen K.F., Lalley P.A.,
RA Weis J.H.;
RA "Isolation of a cDNA that encodes the peptide core of the secretory
RA granule proteoglycan of rat basophilic leukemia-1 cells and
RA assessment of its homology to the human analogue.";
RA J. Biol. Chem. 263:7292-7296(1988).
RN [2]
RX SEQUENCE FROM N.A.
RX TISSUE=Yolk sac carcinoma;
RX MEDLINE=86304425; PubMed=2427521;
RA Bourdon M.A., Shiga M., Ruoslahti E.;
RA "Identification from cDNA of the precursor form of a chondroitin
RA sulfate proteoglycan core protein";
RA J. Biol. Chem. 261:12534-12537(1986).
RN [3]
RX SEQUENCE FROM N.A.
RX MEDLINE=90279692; PubMed=2352541;
RA Giorda R., Chambers W.H., Dahl C.A., Trucco M.;
RA "Isolation and characterization of a cDNA that encodes the core
RA protein of the cytolitic granule proteoglycan in rat natural killer
RA cells.";
RA Nat. Immun. Cell Growth Regul. 9:91-102(1990).
RN [4]
RX SEQUENCE OF 75-179 FROM N.A.
RX MEDLINE=85140287; PubMed=3919394;
RA Bourdon M.A., Oldberg A., Pietschbacher M.D., Ruoslahti E.;
RA "Molecular cloning and sequence analysis of a chondroitin sulfate
RA proteoglycan cDNA.";
RA PROG. Natl. Acad. Sci. U.S.A. 82:1321-1325(1985).

RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RL Nature 368:32-38(1994).
 RN (2)
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SNAP-25 FAMILY.
 CC -!- SIMILARITY: Contains 2 t-SNARE coiled-coil homology domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L14710; AAK84538.1; -;
 DR WormPep; K02D10.5; CE17152.
 DR InterPro; IPR000727; T-SNARE..
 DR SMART; SM00397; t-SNARE; 2.
 DR PROSITE; PS0192; t-SNARE; 2.
 DR Hypothetical protein; Protein transport; Synaptosome; Coiled coil;
 KW Repeat.
 FT DOMAIN 44 106 T-SNARE COILED-COIL HOMOLGY 1.
 FT DOMAIN 179 241 T-SNARE COILED-COIL HOMOLGY 2.
 SQ SEQUENCE 277 AA; 31115 MW; EF33EBAED574B5 CRC64;
 Query Match 11.1%; Score 80; DB 1; Length 277;
 Best Local Similarity 39.7%; Pred. No. 1.5;
 Matches 23; Conservative 10; Mismatches 21; Indels 4; Gaps 2;
 OY 82 PRMDYSLRPYASSSSSSSGSTVSVNQREVFEFFYLDKVLLELDSEERKR 139
 DB 5 PFDDY--RPSAASSTMPVKSYYTTHGYSDEADYYREI--EKLQESLDSTER 58
 RESULT 15
 SLAL_YEAST STANDARD; PRT; 1244 AA.
 AC P32790;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytoskeleton assembly control protein SLAL.
 GN SLAL OR YBL007C OR YBL0321.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RP [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=DDY 228;
 RX MEDLINE=93328765; PubMed=8335689;
 RA Holtzman D.A., Yang S., Drubin D.G.;
 RT "Synthetic-lethal interactions identify two novel genes, SLAL and
 RT SLA2, that control membrane cytoskeleton assembly in Saccharomycetes
 RT cerevisiae."
 RL J. Cell Biol. 122:635-644(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=93070613; PubMed=1441753;
 RA Delaveau T., Jacq C., Perea J.;
 RT "Sequence of a 12.7 kb segment of yeast chromosome II identifies a
 RT PQR-like gene and several new open reading frames."
 RL Yeast 8:761-768(1992).
 CC -!- FUNCTION: ESSENTIAL FOR THE PROPER FORMATION OF THE CORTICAL
 CC ACTIN CYTOSKELETON. INVOLVED IN CONTROLLING THE SIZE OF CORTICAL
 CC PATCHES PERHAPS BY REGULATING THE NUCLEATION OF FILAMENTS AT THE
 CC CORTEX.

CC -!- SIMILARITY: SOME, TO SEA URCHIN BINDIN.
 CC -!- SIMILARITY: Contains 3 SH3 domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z22810; CAA80463.1; -;
 DR EMBL; Z35768; CAA84826.1; -;
 DR EMBL; S47695; AAB23985.1; -;
 DR PIR; S25327; S25327.
 DR HSSP; Q01526; 1AWJ.
 DR SGD; S0000103; SLAL.
 DR GO; GO:0005857; C-actin cortical patch (sensu Saccharomycetes); IDA.
 DR GO; GO:0007015; P-actin filament organization; IMP.
 DR GO; GO:0007047; P-cell wall organization and biogenesis; IMP.
 DR GO; GO:0006897; P-endocytosis; IMP.
 DR GO; GO:0007121; P-polar budding; IMP.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00018; SH3; 3.
 DR Pfam; PF03983; SHD1; 1.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR ProDom; PD000066; SH3; 3.
 DR SMART; SM00326; SH3; 3.
 DR PROSITE; PS50002; SH3; 3.
 KW Cytoskeleton; Actin-binding; SH3 domain; Repeat.
 FT DOMAIN 8 69 SH3 1.
 FT DOMAIN 70 132 SH3 2.
 FT DOMAIN 353 415 SH3 3.
 FT DOMAIN 868 1205
 FT 16 X 7 AA APPROXIMATE REPEATS OF
 FT T-G-G-A-M-M-P.
 FT REPEAT 868 874 1.
 FT REPEAT 877 883 2.
 FT REPEAT 887 893 3.
 FT REPEAT 923 929 4.
 FT REPEAT 945 951 5.
 FT REPEAT 1003 1009 6.
 FT REPEAT 1020 1026 7.
 FT REPEAT 1031 1037 8.
 FT REPEAT 1048 1054 9.
 FT REPEAT 1065 1071 10.
 FT REPEAT 1084 1090 11.
 FT REPEAT 1129 1135 12.
 FT REPEAT 1155 1161 13.
 FT REPEAT 1170 1176 14.
 FT REPEAT 1185 1191 15.
 FT REPEAT 1200 1206 16.
 SQ SEQUENCE 1244 AA; 135948 MW; 7FD85AA776407624 CRC64;
 Query Match 10.9%; Score 78.5; DB 1; Length 1244;
 Best Local Similarity 25.3%; Pred. No. 13;
 Matches 44; Conservative 28; Mismatches 63; Indels 39; Gaps 9;
 OY 3 SNRNN-----QSDDKQARFGVRRPWG-----KPAEIRD-PSRNGALWL- 47
 DB 443 SRSRKSNANASWKDELQNDVVGSAAGKRKSSLSHKKNSATKDFPNKKSRLWVD 502
 OY 48 --GTFETAE-----AARAYDRAAFNGLHAILNFPNE---YPRMDYSLRPPYASS 97
 DB 503 RSGTFKVDAAEFGCAKGIHLHKANGVKIAVADKLSNEDLAYVEKITGFSLE-KFKAND 561
 OY 98 SSSSSGSTTNVSRNQ-----REVFEFYLDKVLLELDSEERKR 139
 DB 562 GSSSRGTDSDSERRRRLKQEKERDRRLKRELKRELKARELLD-EERSR 614
 Search completed: December 17, 2003, 19:21:29
 Job time : 21 secs